PAPER • OPEN ACCESS

Event-by-event study of CR composition with the SPHERE experiment using the 2013 data

View the article online for updates and enhancements.

Event-by-event study of CR composition with the SPHERE experiment using the 2013 data

R.A. Antonov[1]**, T.V. Aulova**[1]**, E.A. Bonvech**[1]**, D.V. Chernov**[1]**, T.A. Dzhatdoev**[1,*]**, Mich. Finger**[2,3]**, Mir. Finger**[2,3]**, V.I. Galkin**[4,1]**, D.A. Podgrudkov**[4,1]**, T.M. Roganova**[1]

[1] Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Leninskie gory 1-2, 119991 Moscow, Russia

[2] Faculty of Mathematics and Physics, Charles University, Ovocn trh 3-5, 116 36 Praha 1, Prague, Czech Republic

[3] Joint Institute for Nuclear Research, Joliot-Curie 6, 141980 Dubna, Moscow region, Russia

[4] Faculty of Physics, Lomonosov Moscow State University, Leninskie gory 1-2, 119991 Moscow, Russia

E-mail: *timur1606@gmail.com

Abstract. We present an event-by-event study of cosmic ray (CR) composition with the reflected Cherenkov light method. The fraction of CR light component above 5 PeV was reconstructed using the 2013 run data of the SPHERE experiment which observed optical Vavilov-Cherenkov radiation of extensive air showers, reflected from snow surface of Lake Baikal. Additionally, we discuss a possibility to improve the elemental groups separability by means of multidimensional criteria.

1. Introduction

The study of the superhigh energy ($E > 10^{15} eV = 1\ PeV$) cosmic ray (CR) composition is the most difficult experimental problem of CR physics. Indeed, even for the most basic primary nuclei mass-dependent quantity, the mean logarithmic mass number $< lnA >$, the spread of results is very high (almost from proton to Iron at some energies) [1]. A somewhat newer review, at the cost of ignoring a part of older works, shows lesser, but still considerable, scatter of $< lnA >$ values [2]. Available information on the elemental groups spectra [3]–[7] at $E > 3$ PeV is almost exclusively due to particle detectors relying on the electron and muon numbers measurement [2]; this method is highly model dependent [8]. Moreover, results obtained with event-by-event techniques are very scarce [9], and most of the above-mentioned works [3]–[7] relied on the deconvolution method, that may suffer from the ill-posedness of the inverse problem [10].

In the present work we describe an event-by-event study of CR composition using the 2013 run data of the SPHERE experiment. The results of this Cherenkov experiment are much less dependent on high-energy hadronic model than for the case of particle detectors and, additionally, are expected to be more stable than ones obtained with the deconvolution method. The datasample used in our analysis is considered in section 2. The method employed for the

composition study is described in section 3. In section 4 the result of the analysis — the fraction of CR light component above 5 PeV — is presented. Finally, in section 5 we show that it is possible to enhance the sensitivity of the method to the primary nuclei mass number with multidimensional criteria.

2. The datasample and low-level data analysis

The SPHERE experiment, stage 2 (2008-2013), employed the SPHERE-2 balloon-borne detector [11] to observe optical Vavilov-Cherenkov radiation ("Cherenkov light") of extensive air showers (EAS), reflected from snow surface of Lake Baikal. The foundations of the method were discussed in [12]. Many details on the detector hardware, observation conditions, simulations, and experimental data analysis could be found in [11]; for brief review see [13]. During the 2013 run, the typical observation altitude (measured by GPS) was in range $H= 400$–$700\ m$ above the snow surface. In total, 3813 events were recorded; 459 of them were recognized as EAS events (for the low-level experimental data analysis methods, see [11], section 5.4). For the next step of the analysis, the lateral distribution function (LDF) reconstruction, we employed a new code, that is generally similar to the code used in our previous works [14], [11], but was written in a completely independent manner. This circumstance allows us to test the robustness of our results against those obtained with the "old" LDF reconstruction method [14].

In total, 421 LDF were obtained with the new method out of the 459 events. The zenith and azimuthal angles of these showers (θ, φ) were measured under an assumption that the shower's front is a plane. Out of the 421 LDF 354 events have estimated zenith angles $\theta < 40°$ and reconstructed axis distance to the center of the detector's field-of-view (FOV) projection to the observation surface $R < R_{Max}$; $R_{Max} = R_{FOV} + 30\ m$, where $R_{FOV} = 0.4903 \cdot H$ is the FOV projection radius. The present analysis relies on a subsample of the late 354 LDF, 328 events. The simulation, however, was perfomed for the full sample of showers with $\theta < 40°$ and $R < R_{Max}$, that corresponds to 354 events. Given that the experimental sample is nearly complete, in what follows we neglect all possible associated biases.

3. The method to separate elemental groups

3.1. The LDF steepness parameter

For composition study we follow the general approach developed in [15]–[16], [14]. Using a sample of model LDF, we define the criteria of the elemental groups separation for different primary energies, zenith angles and observation altitudes. As in [15]–[16], [14], the parameter sensitive to the primary nuclei mass is the LDF "steepness", i.e. the ratio of the number of Cherenkov photons in the circle with the radius of 70 m with the center in the LDF's axis to the same number in the concentric ring with the radii of 70 m and 140 m.

3.2. Simulations

Simulations of EAS development were carried out for primary energies $E_{MC}=$ 10, 30, 100 PeV by means of full direct Monte Carlo (MC) method using the CORSIKA code [17] with the QGSJET-I high energy hadronic model [18] and the GHEISHA low energy hadronic model [19]. The detector response database, that consists of a large number ($\sim 3 \cdot 10^5$) of model showers, was calculated using the Geant4 code [20]. Each "CORSIKA shower" was used multiple times with different axis coordinates. The axis coordinates were uniformly (randomly) distributed over big square with dimensions $1.5 \cdot H \times 1.5 \cdot H$.

As well, the instrumental trigger response was simulated for a range of energies $E_{Trig} =$5-200 PeV, and for a wide range of other conditions. For $E_{Trig} =$5–17.3 PeV, the 10 PeV CORSIKA showers were used, and for $E_{Trig} >$17.3 PeV — the 30 PeV showers. The model showers with $E_{Trig}/E_{MC} = K$ were obtained by multipying the corresponding responses to the factor K. We
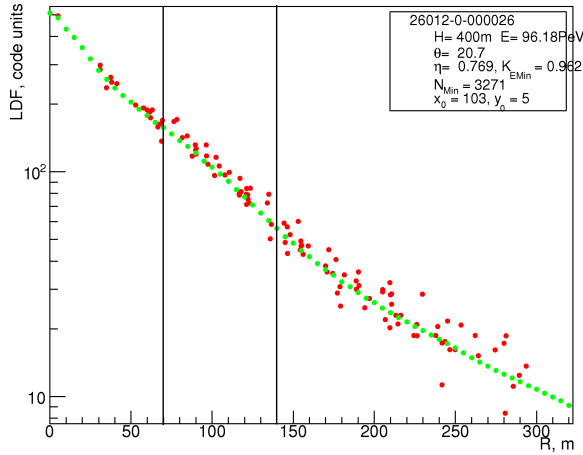
Figure 1. An example of the individual model LDF (red circles) together with a "composite model LDF" (green circles) that fits the individual LDF. The energy of primary nucleus was set to 100 PeV. The legend shows the values of some "true MC" parameters; (x_0, y_0) $[m]$ is the axis position coordinates with respect to the center of the detector's FOV projection to the observation surface. The zenith angle value θ is measured in $°$. As well, the reconstructed energy value (96.18 PeV) is shown.
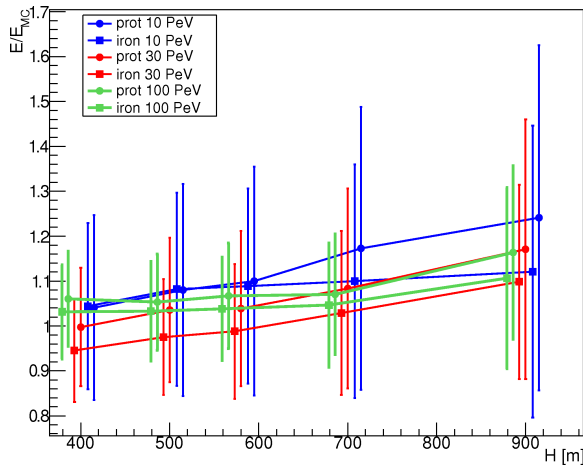

Figure 2. Reconstructed energy for model showers with various primary nuclei, primary energies and observation altitudes.
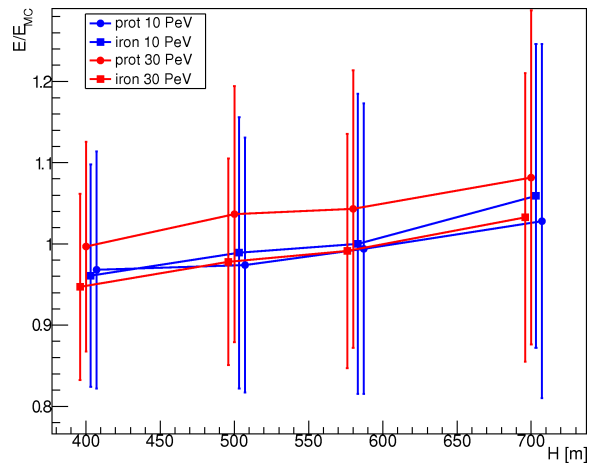

Figure 3. The same, as in figure 2, but only for showers that were registered by the model of the instrumental trigger.

have verified that the E_{MC} values discreteness has a negligible impact on results of the trigger response simulation.

3.3. Estimation of shower parameters

As a first step of the elemental groups separation procedure, we have estimated the LDF steepness parameter η for all model showers, integrated on time, by fitting these "discrete LDF" by a set of specially calculated smooth continious curves. Each curve of the set (we call it a "composite model LDF", CLDF) was composed from many model "discrete LDF" corresponding to the same CORSIKA shower. An example of a CLDF is shown in figure 1 by green circles, and the corresponding individual model LDF is plotted by red circles. The normalisation factor of a CLDF, that is nearly proportional to the primary energy [21], was utilized as the primary energy estimator, and the CLDF steepness parameter — as an estimate of the η value. The regions (0, 70) m and (70, 140) m used for the η parameter estimation are constrained by vertical black lines in figure 1.

The energy reconstruction performances of the described procedure are presented in figures 2-3. Figure 2 deals with all showers in the model sample, irrespectively of the axis position. The mean reconstructed energy values are shown by symbols for proton and Iron primaries for the
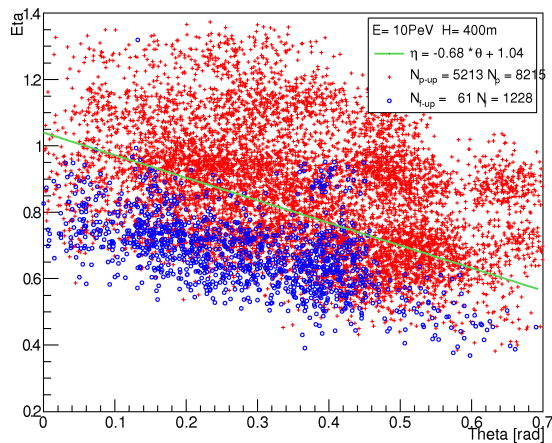
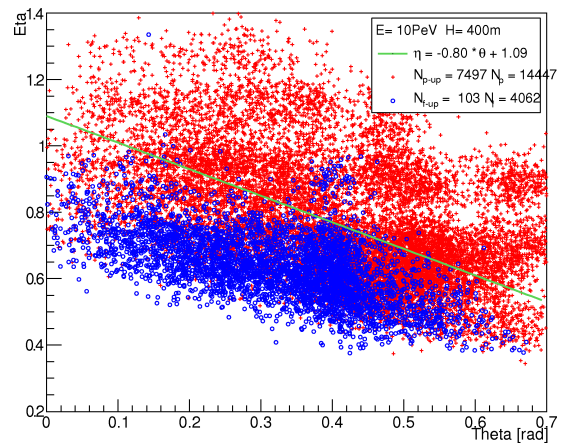**Figure 4.** The result of the model showers classification for E_{Trig}= 11.2 PeV.

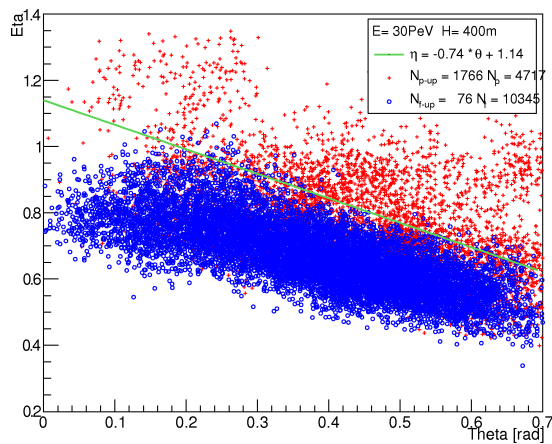**Figure 5.** The same, as in figure 4, but for E_{Trig}= 15.9 PeV.

Figure 6. The same, as in figures 4-5, but for E_{Trig}= 20.9 PeV.
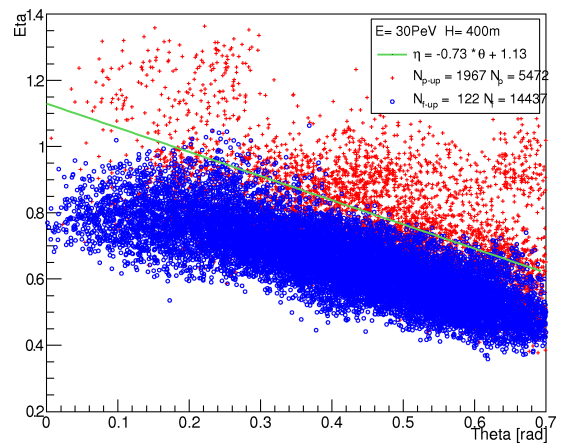

Figure 7. The same, as in figures 4-6, but for E_{Trig}= 29.6 PeV.
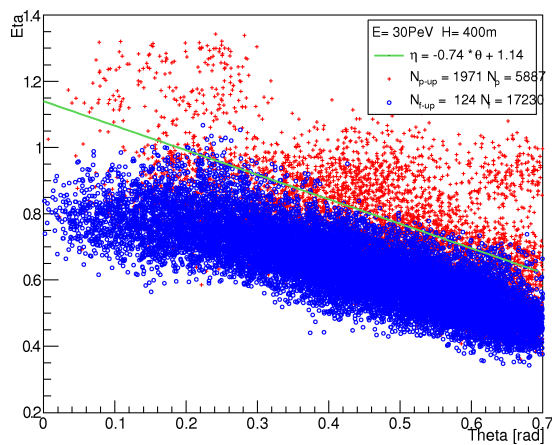

Figure 8. The same, as in figures 4-7, but for E_{Trig}= 41.8 PeV.

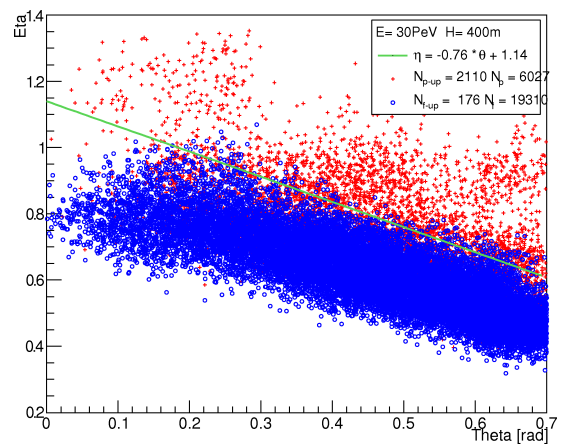**Figure 9.** The same, as in figures 4-8, but for E_{Trig}= 102.6 PeV.

primary ("true MC") energy values E_{MC}= (10, 30, 100) PeV, and several altitudes (400, 500, 580, 700, 900) m. The corresponding standard deviations σ_E (σ_E is a measure of the primary energy statistical reconstruction uncertainty) are shown by bars. A small artificial shift between the results for different conditions is introduced on the H axis to make them visible. Similar results for E_{MC}= 10 PeV and 30 PeV are shown in figure 3 for showers that satisfy the trigger condition. We do not show results for 900 m, because the maximal altitude for the 2013 run was ≈700 m. It is evident that the statistical uncertainty of the primary energy value is qiute low (from ~10 % to ~20 % depending from the primary energy and the altitude values), and the systematic error is moderate (<5 % for most of the cases).

3.4. Separation of the elemental groups

To study the CR composition, we employed a set of Bayesian classifiers [22] using the above-described model LDF sample as a training set. In the present work, for simplicity, we constrain ourselves to the case of the linear border between the classes of primary nuclei, and two event classes — proton and Iron primary nuclei. Following the famous "Bayes rule", we have set the prior probabilities 0.5 for both classes. The results of the classifier training are shown for the altitude value H= 400 m and different energies E_{Trig} in figures 4-9. The values of the LDF steepness parameter for proton (red crosses) and Iron (blue circles) vs. the primary zenith angle value θ are plotted; the boundary between the classes is shown by green straight line. The E_{MC} parameter value is shown (see subsection 3.1 for explanation), as well as the parameters of the border line equation. The numbers of proton (N_{p-up} and N_p) and Iron (N_{f-up} and N_f) showers above the border, and total numbers of these model showers, respectively, are presented. To suppress statistical fluctuations arising from a small number of CORSIKA showers, we have introduced a small Gaussian smearing on both parameters (θ, η).

The fraction of model proton showers classified as light nuclei is $F_p = N_{p-up}/N_p$= (0.635, 0.519, 0.374, 0.359, 0.335, 0.350) for the E_{Trig} values (11.2, 15.9, 20.9, 29.6, 41.8, 102.6) PeV, respectively, and the corresponding fraction of Iron contamination $F_f = N_{f-up}/N_f$= (0.050, 0.025, 0.007, 0.008, 0.007, 0.009). It is evident from figures 4-5 that Iron nuclei are effectively suppressed by the instrumental trigger effects; at sufficiently low energy the SPHERE-2 detector could observe practically only light nuclei. Overall, the Iron contamination appears to be quite low. Results such as shown in figures 4-9 were obtained for 5 altitudes (400, 500, 580, 700, 900) m and 54 logarithmically spaced energy bins from 5 to 200 PeV.

All experimental showers from the sample defined in section 2 have the parameters (θ, E, η) estimated, and the altitude value H measured. These events were classified as proton-like or Iron-like according to their position above or below the border. The total number of the light CR component showers in 6 energy bins in the energy range 5-200 PeV was reconstructed as $N_{i-light} = N_{i-p-like}/F_p$, where $N_{i-p-like}$ is the number of experimental showers above the border in the ith bin. The light and heavy component intensities were corrected for the instrumental acceptance effects to account for the difference between the detection efficiencies of light and heavy nuclei.

In comparison to [14], in the present work we were able to substantially lower the threshold of the method using calculations of the instrumental acceptance [23],[11]. The first-order correction in the present work uses the 2013-integrated acceptance curves for the case of $R < R_{FOV} + 100$ m, instead of $R < R_{FOV} + 30$ m. The number of available model showers at E= 100 PeV is currently low, therefore the analysis was performed with the 10 PeV and 30 PeV model showers. Corrections for these inaccuracies are made in section 4.

4. Results

The result of our analysis — the estimated fraction of light nuclei in the energy range 5-200 PeV — is shown in figure 10 by red stars with statistical uncertainties. For the case of the
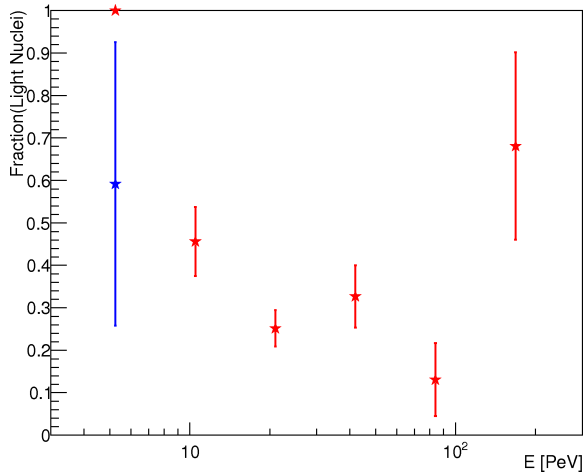
Figure 10. The fraction of light ("proton-like") nuclei reconstructed using the 2013 run data of the SPHERE experiment.
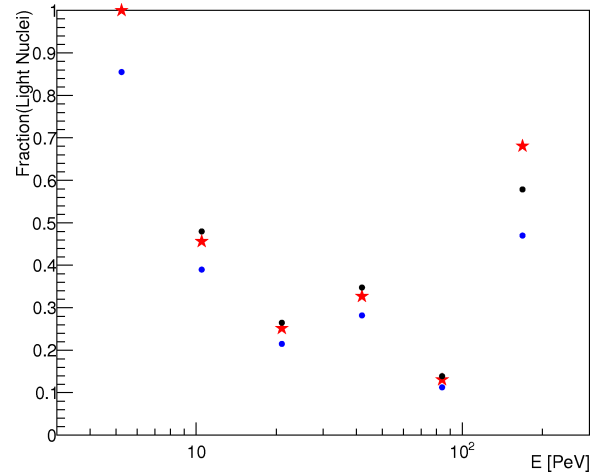

Figure 11. The fraction of light nuclei before (red stars) and after corrections (black and blue circles; see text for more details).

120–240 PeV energy interval the statistical uncertainty is quite high and our results are not constraining. For the case of 3.5–7 PeV energy interval we didn't find any indication of the Iron nuclei presence. To estimate the statistical uncertainty for this bin, we have added one "bonus" Iron shower; the corresponding result is shown in figure 10 by the blue star. The light nuclei fraction behaviour presented in figure 10 is qualitatively similar to the KASCADE-Grande result [7], and doesn't contradict our earlier estimate [14] obtained with the 2012 run data. We have already applied the correction for model energy discreteness here (see section 3).

Several systematic effects do exist that could change our results. A shift of the light nuclei fraction could occur as a result of the energy estimation systematic uncertainty (see figure 3). For some energy bins, this effect tends to enhance the reconstructed light nuclei intensity up to 10–15 % with respect to its true value. Therefore, the corrected fraction of light nuclei in these bins would be lower than plotted in figure 10 (for instance, 25-27 % instead of 30 %). We have estimated and applied corrections for this effect, as well as the acceptance correction discussed at the end of section 3. The new result after these corrections is shown in figure 11 by black circles. As well, we have estimated the LDF reconstruction procedure systematic uncertainty, and the corresponding result including this last correction, as well as all previous ones, is shown in figure 11 by blue circles. Dedicated work to study these and other systematic effects is in progress, and will be reported elsewhere.

5. Multidimensional method for superhigh energy CR composition study

In sections 3–4 were have employed the set of one-dimensional criteria, relying on the $\eta = \eta(A, E_{Trig}, \theta, H)$ parameter. Here we show that it is possible to improve the nuclei classes separability by using multidimensional pattern recognition methods. We demonstrate the case on our model sample (without imposing the trigger effects) by studiyng classification performances for 3 nuclei groups separation — proton, Nitrogen, and Iron. The feature vector was chosen to be 4-dimensional, and was composed of Cherenkov light intensities in concentric rings with the center in the shower's axis and with radii (0, 40) m, (40, 80) m, (80, 120) m, and (120, 160) m. The zenith angle value θ was added as the fifth classification parameter. The multivariate Bayesian pattern recognition technique under assumption of multidimensional Gaussian distribution of features [22] was used for this study. This technique was recently shown to be very effective for the case of CR composition study with lateral-angular distribution of

Cherenkov light [24]–[25].

5.1. Proton–Nitrogen separation

All graphs in this subsection (figures 12-15) show contamination of Nitrogen nuclei (i.e. the fraction of Nitrogen nuclei classified as protons) vs. proton selection efficiency. Black curves are drawn for $H= 400\ m$, red curves — for $H= 580\ m$, and blue curves — for $H= 900\ m$. The primary energy values and zenith angle range are shown in the captions to the graphs.
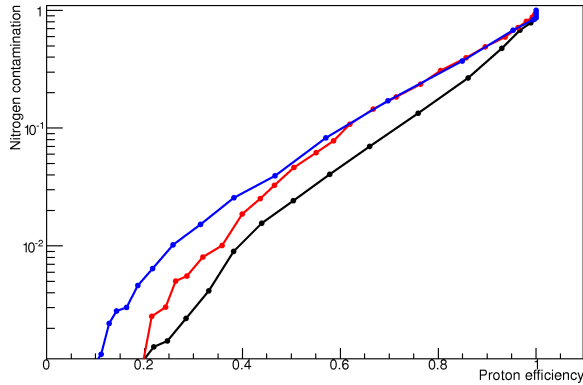

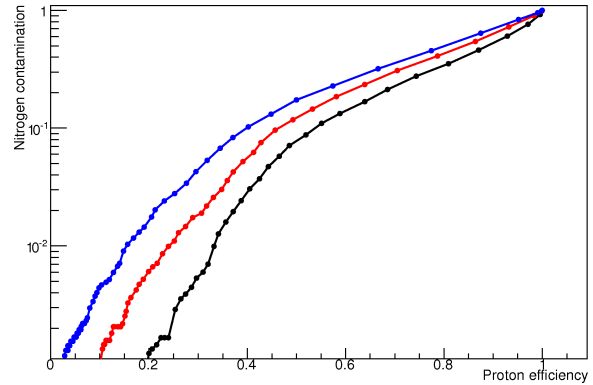**Figure 12.** Proton-Nitrogen separation for $E= 10\ PeV$ and $\theta= 0\text{-}20°$.

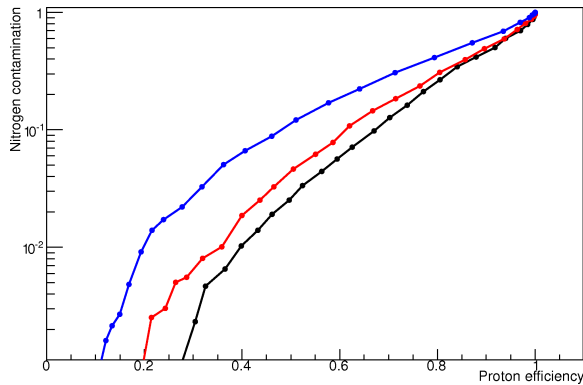**Figure 13.** The same, as in figure 12, but for $\theta= 20\text{-}40°$.

Figure 14. The same, as in figures 12-13, but for $E= 30\ PeV$ and $\theta= 0\text{-}20°$.
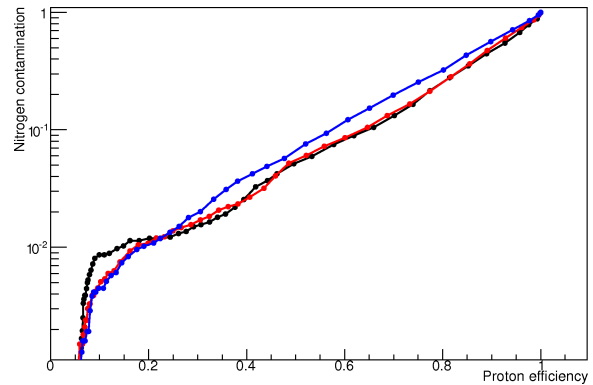

Figure 15. The same, as in figure 14, but for $\theta= 20\text{-}40°$.

5.2. Proton–Iron, proton–Helium and Nitrogen–Iron separation

The same results as in figures 12-15, but for separation of other elemental groups, were also obtained. For the case of proton-Iron separation (except the 10 PeV and $\theta= 20\text{-}40°$ case), the typical proton selection efficiency corresponding to $\sim$1 % Iron contamination is $\approx$60-90 %, depending on conditions. For the 10 PeV and $\theta= 20\text{-}40°$ case, the proton selection efficiency is lower, but in this case the Iron showers would be suppressed by the trigger effects, as was explained in section 3. The same applies to other cases where proton selection efficiency is <75 %; therefore, the proton selection efficiency is, in fact, >75 % for $\sim$1 % Iron contamination.

As well, we have studied the Nitrogen–Iron classification performance, and obtained the typical Nitrogen selection efficiency value $\approx$40 % for $\sim$1 % Iron contamination. To conclude,

let us note that proton–Helium separation is, in principle, also possible with the presented multidimensional technique. However, with the criteria considered in the present work, the proton selection efficiency was observed to be always <25 % for ∼1 % Helium contamination.

6. Conclusions

We have presented and discussed the method for event-by-event study of superhigh CR composition applicable to the SPHERE experiment working conditions. It was found that a part of the light component showers can be readily selected against the heavy component background. The primary energy can be reconstructed with sufficiently low uncertainty, both statistical and systematic. The result of the analysis — the light nuclei fraction vs. energy — is qualitatively similar to the KASCADE-Grande result, and doesn't contradict to our earlier estimates. As well, we have shown that multidimensional criteria, in principle, allow to enhance the separability of the elemental group classes.

References

[1] Tsunesada Y et al. (BASJE) 2008 *Proc. Int. Conf. on Cosmic Rays* vol 4 127
[2] Kampert K-H and Unger M 2012 *APh* **35** 660
[3] Antoni T et al. (KASCADE) 2005 *APh* **24** 1
[4] Antoni T et al. (KASCADE) 2009 *APh* **31** 86
[5] Garyaka A et al. (GAMMA) 2007 *APh* **28** 169
[6] Apel W D et al. (KASCADE-Grande) 2013 *APh* **47** 54
[7] Apel W D et al. (KASCADE-Grande) 2013 *Phys. Rev.* **D87** 081101(R)
[8] Apel W D et al. (KASCADE-Grande) 2014 *Advances in Space Research* **53** 1456
[9] Antoni T et al. (KASCADE-Grande) 2002 *APh* **16** 245
[10] Ter-Antonyan S V 2007 *APh* **28** 321
[11] Antonov R A and et al. 2015 *Physics of Particles and Nuclei* **46** 60
[12] Chudakov A E 1974 *Proc. All-USSR Symp. on Exp. Meth. of UHECR, Yakutsk* (Yakutsk: Russian Academy of Sciences Publisher) (In Russian) 69
[13] Antonov R A and et al. 2015 *DESY Conf. Proc. (PANIC-2014)* (In print)
[14] Antonov R A and et al. 2013 *J. Phys.: Conf. Series* **409** 012088
[15] Anokhina A M and et al. 2009 *Bulletin of the Lebedev Physics Institute* **36** 146
[16] Antonov R A and et al. 2009 *Proc. Int. Conf. on Cosmic Rays (Lodz)* vol HE.1.3 (Lodz:) id. 434
[17] Heck D and et al. 1998 *FZKA Report* **6019**
[18] Kalmykov N et al. 1997 *Nucl. Phys. Proc. Suppl.* B **52** 17
[19] Fesefeldt H C 1985 *Technical Report* **PITHA 85-02 RWTH**
[20] Agostinelli S and et al. 2003 *NIM* A **506** 250
[21] Dedenko L G and et al. 2004 *Nucl. Phys.: Proc. Suppl.* B **136** 1217
[22] Theodoridis S and Koutroumbas K 2003 *Pattern Recognition* 2nd edition (Amsterdam: Elsevier)
[23] Antonov R A and et al. 2013 *Proc. Int. Conf. on Cosmic Rays (Rio de Janeiro)* (Rio de Janeiro:) id. 1185
[24] Galkin V I and Dzhatdoev T A 2011 *Izvestiya Rossiiskoi Akademii Nauk* **75** 338 (In Russian)
[25] Borisov A S and Galkin V I 2013 *J. Phys.: Conf. Series* **409** 012089